



Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics  
(IMACM)

Preprint BUW-IMACM 19/33

Christoph Hachtel, Andreas Bartel, Michael Günther, Adrian Sandu

**Multirate Implicit Euler Schemes for a Class of  
Differential-Algebraic Equations of Index-1**

December 3, 2019

<http://www.math.uni-wuppertal.de>

# Multirate Implicit Euler Schemes for a Class of Differential-Algebraic Equations of Index-1

Christoph Hachtel<sup>a,\*</sup>, Andreas Bartel<sup>a</sup>, Michael Günther<sup>a</sup>, Adrian Sandu<sup>b</sup>

<sup>a</sup>*Bergische Universität Wuppertal, Chair of Applied Mathematics and Numerical Analysis (AMNA), Gaußstraße 20, 42119 Wuppertal, Germany.*

<sup>b</sup>*Virginia Polytechnic Institute and State University, Computational Science Laboratory, Department of Computer Science, 2202 Kraft Drive, Blacksburg, VA 22060, USA*

---

## Abstract

Systems of differential equations which consist of subsystems with widely different dynamical behaviour can be integrated by multirate time integration schemes to increase the efficiency. These schemes allow the usage of inherent step sizes according to the dynamical properties of the subsystem. In this paper, we extend the multirate implicit Euler method to semi-explicit differential-algebraic equations of index-1 where the algebraic constraints only occur in the slow changing subsystem. We discuss different coupling approaches and show that consistency and convergence order 1 can be reached. Numerical experiments validate the analytical results.

*Keywords:* Numerical Analysis, Differential Algebraic Equations, Multirate Time Integration

---

## 1. Motivation and Introduction

The mathematical modelling of a technical or physical problem often leads to a system of differential equations, where different components provide a widely spread dynamical behaviour: some components of the system are changing much faster than others. In time domain simulation, the fastest component will define the maximal step size of the integration scheme. Therefore the simulation of the complete system will be very costly and inefficient. To overcome this problem, a multirate time integration scheme can be applied. These schemes use inherent step sizes for the components according to their dynamical behaviour: fast components are integrated with a small step size and slower ones with a larger step size. The crucial part is the coupling between the components: how can the values of the fast components be approximated during the integration of the slow components and vice versa. For systems of ordinary differential equations (ODEs) multirate schemes with different coupling approaches have been widely discussed in e.g. [1, 2, 3, 4, 5].

In many applications (e.g. circuit simulation, electro-magnetic field simulation), the system cannot be modelled as a system of ODEs since algebraic constraints occur due to an automated modelling approach. Such systems are called differential-algebraic equations (DAEs). The numerical treatment of DAEs differs from the classical ODE time integration framework and is already for a global time stepping (single-rate) method more challenging [6, 7, 8].

---

\*Corresponding author

*Email addresses:* [hachtel@math.uni-wuppertal.de](mailto:hachtel@math.uni-wuppertal.de) (Christoph Hachtel), [bartel@math.uni-wuppertal.de](mailto:bartel@math.uni-wuppertal.de) (Andreas Bartel), [guenther@math.uni-wuppertal.de](mailto:guenther@math.uni-wuppertal.de) (Michael Günther), [asandu@cs.vt.edu](mailto:asandu@cs.vt.edu) (Adrian Sandu)

Multirate time integration for DAEs has been discussed in [9] for mixed-multirate methods based on ROW schemes. In [10] multirate schemes for DAEs based on BDF-methods are presented using a specialized stability analysis. In this work, we extend the multirate implicit Euler method to semi-explicit DAEs of index-1, where the algebraic constraints only occur in the slow changing variables. We present different coupling strategies between the subsystems: *Decoupled-Slowest-First* analogue to the slow-fast method in [10], *Coupled-Slowest-First* which is similar to compound-fast in [10] and *Coupled-First-Step* which coincides with the mixed-multirate approach in [9]. We provide a direct convergence analysis and show that convergence order 1 can be reached. Numerical results verify the theoretical results. The methods applied can be extended to higher-order one-step methods which will be part of future work.

The paper is organised as follows: In Section 2, we briefly introduce the multirate implicit Euler method for ODEs and extend the scheme to DAEs. Thereby we discuss three different coupling approaches. A detailed consistency analysis for all coupling approaches follows (Section 3). The convergence of the schemes is addressed in Section 4 and numerical results are presented in Section 5.

## 2. Multirate Implicit Euler Method

After introducing the multirate implicit Euler-method for ODEs, we extend the integration method to semi-explicit DAEs of index-1 and present three different coupling approaches.

### 2.1. Multirate Implicit Euler Method: ODE-Case

We consider the set of coupled initial value problems (IVP) of ODEs:

$$\dot{y}_F(t) = f_F(t, y_F, y_S), \quad y_F(t_0) = y_{F,0}, \quad (1)$$

$$\dot{y}_S(t) = f_S(t, y_S, y_S), \quad y_S(t_0) = y_{S,0}, \quad (2)$$

for time  $t \in [t_0, t_{\text{end}}]$ , a fast ( $F$ ) changing variable  $y_F(t) \in \mathbb{R}^{n_F}$  and a slow ( $S$ ) changing variable  $y_S(t) \in \mathbb{R}^{n_S}$ . The multirate implicit Euler method (mrIRK1 for **m**ultirate **I**mplicit **R**unge-**K**utta method of order **1**) integrates the fast changing subsystem (1) with a small micro-step size  $h$  and the slow changing subsystem (2) with a large macro-step size  $H$ . We assume a fixed multirate factor, i.e.,  $m = H/h$  for an integer  $m \in \mathbb{N}$ . Given approximations  $y_{F,n}, y_{S,n}$  at  $t_n$ , the integration over the macro-step  $t_n \rightarrow t_{n+1} = t_n + H$  with micro grid points  $t_{n+l/m} = t_n + lh$  ( $l = 0, 1, \dots, m-1$ ) reads:

$$y_{F,n+(l+1)/m} = y_{F,n+l/m} + h f_F(t_{n+(l+1)/m}, y_{F,n+(l+1)/m}, \bar{y}_{S,n+(l+1)/m}) \quad (l=0,1,\dots,m-1), \quad (3)$$

$$y_{S,n+1} = y_{S,n} + H f_S(t_{n+1}, \bar{y}_{F,n+1}, y_{S,n+1}),$$

where  $\bar{y}_F, \bar{y}_S$  denote the coupling variables to the other subsystem. There are different strategies to define the values of the coupling variables. We discuss important coupling terms after the extension to DAEs.

### 2.2. Multirate Implicit Euler Method: DAE-Case

We consider the following index-1 DAE-IVP in semi-explicit form

$$\begin{aligned} \dot{y}_F(t) &= f_F(y_F, y_S, z_S), & y_F(t_0) &= y_{F,0}, \\ \dot{y}_S(t) &= f_S(y_F, y_S, z_S), & y_S(t_0) &= y_{S,0}, \\ 0 &= g_S(y_F, y_S, z_S), & z_S(t_0) &= z_{S,0} \end{aligned} \quad (4)$$

for  $t \in [t_0, t_{\text{end}}]$ , fast differential variable  $y_F(t) \in \mathbb{R}^{n_F}$ , slow differential variable  $y_S(t) \in \mathbb{R}^{n_S}$  and slow algebraic variable  $z_S(t) \in \mathbb{R}^{n_Z}$ . The initial values (IVs) shall consistent and index-1 be guaranteed by the assumption

$$\det \left( \frac{\partial g_S}{\partial z_S} \right) \neq 0$$

in a neighbourhood of the analytic solution. We point out that the algebraic constraints only occur in the slow subsystem but may depend on the solution of the fast subsystem. Such a coupling structure may arise field/circuit problems, where a slowly changing electro-magnetic field is coupled with fast changing, regularised electric circuit model.

By the index-1 condition, we can solve the algebraic constraint (locally) for the algebraic variable  $z_S$  using the implicit function theorem:

$$z_S = G(y_F, y_S). \quad (5)$$

A DAE of index-1 can be integrated with an implicit Euler method in single-rate. Thereby, the ODE convergence properties are be maintained, since the method is stiffly accurate [6].

To exploit the multirate behaviour of the DAE (4), we propose the following integration method based on the mrIRK1 method (3) and the classical single-rate implicit Euler method for DAEs of index-1 [6]. The integration of system (4) over the macro-step  $t_n \rightarrow t_{n+1} = t_n + H$  reads:

$$y_{F,n+(l+1)/m} = y_{F,n+l/m} + h f_F(y_{F,n+(l+1)/m}, \bar{y}_{S,n+(l+1)/m}, \bar{z}_{S,n+(l+1)/m}) \quad (6)$$

$(l=0,1,\dots,m-1)$

$$y_{S,n+1} = y_{S,n} + H f_S(\bar{y}_{F,n+1}, y_{S,n+1}, z_{S,n+1}) \quad (7)$$

$$0 = g_S(\bar{y}_{F,n+1}, y_{S,n+1}, z_{S,n+1}). \quad (8)$$

The coupling variables are denoted by  $\bar{y}_F, \bar{y}_S, \bar{z}_S$ . We refer to this method as mrIRK1-DAE.

### 2.3. Multirate Implicit Euler Method: Coupling Strategies

We briefly introduce three different strategies how the coupling terms can be realised (on the macro step  $t_n \rightarrow t_n + H$ ).

*Decoupled-Slowest-First.* For an approximation at  $t_n + H$ , the slow subsystem is first solved via (7-8). Thereby, the coupling variable  $\bar{y}_F$  is fixed by constant extrapolation  $\bar{y}_{F,n+1} = y_{F,n}$ . Then, for the integration of the fast subsystem via (6), the slow differential coupling variable  $\bar{y}_S$  is linearly interpolated on the micro-step level:

$$\bar{y}_{S,n+l/m} = \frac{m-l}{m} y_{S,n} + \frac{l}{m} y_{S,n+1}. \quad (9)$$

Analogously, the algebraic coupling variable  $\bar{z}_{S,n+l/m}$  can be interpolated. Another approach for  $\bar{z}_S$  is the implicit definition via the non-linear equation

$$\begin{aligned} y_{F,n+(l+1)/m} &= y_{F,n+l/m} + h f_F(y_{F,n+(l+1)/m}, \bar{y}_{S,n+(l+1)/m}, \bar{z}_{S,n+(l+1)/m}), \\ 0 &= g(y_{F,n+(l+1)/m}, \bar{y}_{S,n+(l+1)/m}, \bar{z}_{S,n+(l+1)/m}). \end{aligned} \quad (10)$$

We show that both realisations of the algebraic-to-fast coupling lead to the same consistency order of the integration method.

*Coupled-Fastest-First.* We remark that starting the computation with the fast subsystem and extrapolating the slow variables is conceivable. However, the slowest-first approach fits better to a step-size control on the macro-step level and, thus, is more relevant for practical applications [1]. Therefore, this strategy is not further address here.

*Coupled-Slowest-First.* This approach was introduced in [4] based on a  $\theta$ -method. The idea is the following: the complete system (4) is solved on the macro-step level

$$\begin{aligned} y_{F,n+1}^* &= y_{F,n} + H \cdot f_F(y_{F,n+1}^*, y_{S,n+1}, z_{S,n+1}) \\ y_{S,n+1} &= y_{S,n} + H \cdot f_S(y_{F,n+1}^*, y_{S,n+1}, z_{S,n+1}) \\ 0 &= g_S(y_{F,n+1}^*, y_{S,n+1}, z_{S,n+1}). \end{aligned} \quad (11)$$

The step-size  $H$  is chosen according to the dynamical properties of  $y_S$  and  $z_S$ . Thus, the approximation  $y_{F,n+1}^*$  is not accurate and therefore refused. The integration of the fast subsystem is re-computed using micro-steps (6) and coupling variables  $\bar{y}_S$ ,  $\bar{z}_S$  are linearly interpolated (9). Alternatively,  $\bar{z}_S$  can be computed via the non-linear algebraic constraint, see (10).

*Coupled-First-Step.* Here, the first micro-step of the fast subsystem is computed together with macro-step of the slow subsystem. This technique was introduced for Runge-Kutta based schemes in [3]. The compound-step reads:

$$\begin{aligned} y_{F,n+1/m} &= y_{F,n} + h f_F(y_{F,n+1/m}, y_{S,n+1}, z_{S,n+1}) \\ y_{S,n+1} &= y_{S,n} + H f_S(y_{F,n+1/m}, y_{S,n+1}, z_{S,n+1}) \\ 0 &= g_S(y_{F,n+1/m}, y_{S,n+1}, z_{S,n+1}). \end{aligned} \quad (12)$$

The remaining micro-steps of the fast subsystem are computed according to (6) for  $l = 1, \dots, m-1$ . Here, the slow coupling variables  $\bar{y}_S$ ,  $\bar{z}_S$  can be obtained by linear interpolation (9) or, alternatively,  $\bar{z}_S$  via the algebraic constraint (10).

### 3. Consistency Analysis

We estimate the error that is made during one macro-step  $t_n \rightarrow t_{n+1} = t_n + H$  caused by the mrIRK1-DAE method (based on  $m$  micro-steps; i.e.,  $H = m \cdot h$ ). We discuss the three introduced coupling strategies.

#### 3.1. Preliminaries

Let  $x : [t_0, t_{\text{end}}] \rightarrow \mathbb{R}^k$  denote some set of variables (of the above DAE) and let exact initial values  $x(t_n)$  be given for the macro-step  $[t_n, t_{n+1}]$ . At the end of the macro-step ( $t = t_{n+1}$ ), we have a numerical approximation  $x_{n+1}$  of an analytic solution  $x(t_{n+1})$  and the error notation:

$$\Delta x_{n+1} := x_{n+1} - x(t_{n+1}). \quad (13)$$

Hence, we assume at  $t = t_n$ : (for any vector norm  $\|\cdot\|$ )

$$\|\Delta y_{F,n}\| = \|\Delta y_{S,n}\| = \|\Delta z_{S,n}\| = 0. \quad (14)$$

For simplicity of notation, we introduce the following sloppy short-hand on the  $n$ th macro-step:

$$\|x(t)\|_\infty := \max_{\tau \in [t_n, t_{n+1}]} \|x(\tau)\|.$$

The following assumption is valid for the whole section.

**Assumption 1.** For some  $\varepsilon > 0$  and the analytic solution  $(y_F(\cdot), y_S(\cdot), z_S(\cdot))$  of the DAE (4), we define the neighbourhood at time  $\tau$

$$\mathcal{E}(\tau) := \{(y_F, y_S, z_S) \in \mathbb{R}^{n_F+n_S+n_Z} \mid \|y_F - y_F(\tau)\|, \|y_S - y_S(\tau)\|, \|z_S - z_S(\tau)\| \leq \varepsilon\}$$

and assume the following:

- (i) The right-hand sides of DAE (4)  $f_F, f_S, g_S$  are sufficiently smooth and all first and second partial derivatives are (locally) uniformly bounded. The Lipschitz constant of  $f_F$  with respect to  $y_S$  reads

$$L_{FS} := \max_{\tau \in [t_n, t_{n+1}], \mathcal{E}(\tau)} \left\| \frac{\partial f_F}{\partial y_S}(y_F, y_S, z_S) \right\|, \quad (15)$$

and  $L_{FF}, L_{FZ}, L_{SF}, L_{SS}, L_{SZ}$  are defined analogously.

- (ii) For DAE (4), the implicit function  $G$  (5) shall exist globally on  $[t_n, t_{n+1}]$ .  $G$  shall be sufficiently smooth and the partial derivatives shall be uniformly bounded. The corresponding Lipschitz constant reads:

$$L_{GS} := \max_{\tau \in [t_n, t_{n+1}], \mathcal{E}(\tau)} \left\| \frac{\partial G}{\partial y_S}(y_F, y_S) \right\| \quad (16)$$

and  $L_{GF}$  analogously.

Except for the first step in the coupled-first-step strategy, the computation of the fast components is the same. Thus, we start the error estimation for the fast subsystem.

### 3.2. Accuracy of the Fast Components

For a macro-step  $[t_n, t_{n+1}]$ , we have:

**Lemma 1.** Let be given an index-1 DAE-IVP on  $[t_n, t_{n+1}]$  (4), which fulfils Ass. 1. Let the approximation  $y_{F,n+1}, y_{S,n+1}, z_{S,n+1}$  be computed by the *mrIRK1-DAE* scheme (6-8) with macro-step size  $H$  and micro-step size  $h = H/m$  ( $m \in \mathbb{N}$ ). If the micro-step size is restricted to  $0 < 1 - hL_{FF} < 1$ , then the error in the fast subsystem after one macro-step  $t_n \rightarrow t_n + H$  can be bounded by

$$\begin{aligned} \|\Delta y_{F,n+1}\| &\leq C \left[ \frac{H^2}{2m} \|\ddot{y}_F(t)\|_\infty \right. \\ &\quad \left. + h \sum_{k=0}^{m-1} (L_{FS} \|\Delta \bar{y}_{S,n+(k+1)/m}\| + L_{FZ} \|\Delta \bar{z}_{S,n+(k+1)/m}\|) \right] \end{aligned} \quad (17)$$

with a constant  $C > \left(\frac{1}{1-hL_{FF}}\right)^m > 0$  and the coupling errors  $\Delta \bar{y}_{S,n+(k+1)/m}, \Delta \bar{z}_{S,n+(k+1)/m}$ .

*Proof.* We estimate  $\Delta y_{F,n+(l+1)/m}$  in one micro-step  $t_{n+l/m} \rightarrow t_{n+(l+1)/m}$ :

$$\begin{aligned} \Delta y_{F,n+(l+1)/m} &= \underbrace{y_{F,n+l/m} - y_F(t_{n+l/m})}_{=\Delta y_{F,n+l/m}} + (-y_F(t_{n+(l+1)/m}) + y_F(t_{n+l/m})) \\ &\quad + hf_F(y_F(t_{n+(l+1)/m}), y_S(t_{n+(l+1)/m}), z_S(t_{n+(l+1)/m})) \\ &\quad + hf_F(y_{F,n+(l+1)/m}, \bar{y}_{S,n+(l+1)/m}, \bar{z}_{S,n+(l+1)/m}) \\ &\quad - hf_F(y_F(t_{n+(l+1)/m}), y_S(t_{n+(l+1)/m}), z_S(t_{n+(l+1)/m})). \end{aligned}$$

The local truncation error of the single-rate implicit Euler method is defined as

$$\delta_{n+l/m} = y_F(t_{n+l/m}) + hf_F(y_F(t_{n+(l+1)/m}), y_S(t_{n+(l+1)/m}), z_S(t_{n+(l+1)/m})) - y_F(t_{n+(l+1)/m}).$$

Applying the mean-value theorem, we get

$$\begin{aligned} \Delta y_{F,n+(l+1)/m} &= \Delta y_{F,n+l/m} + \delta_{n+l/m} + h \int_0^1 \frac{\partial f_F}{\partial y_F}(\Theta(\sigma)) \Delta y_{F,n+(l+1)/m} d\sigma \\ &+ h \int_0^1 \frac{\partial f_F}{\partial y_S}(\Theta(\sigma)) \Delta \bar{y}_{S,n+(l+1)/m} d\sigma + h \int_0^1 \frac{\partial f_F}{\partial z_S}(\Theta(\sigma)) \Delta \bar{z}_{S,n+(l+1)/m} d\sigma \end{aligned}$$

with evaluation at

$$\Theta(\sigma) := \begin{pmatrix} y_F(t_{n+(l+1)/m}) + \sigma \Delta y_{F,n+(l+1)/m} \\ y_S(t_{n+(l+1)/m}) + \sigma \Delta \bar{y}_{S,n+(l+1)/m} \\ z_S(t_{n+(l+1)/m}) + \sigma \Delta \bar{z}_{S,n+(l+1)/m} \end{pmatrix}.$$

Applying norms and using Lipschitz continuity, we can estimate:

$$\begin{aligned} \|\Delta y_{F,n+(l+1)/m}\| &\leq \|\Delta y_{F,n+l/m}\| + \frac{h^2}{2} \|\ddot{y}_F(t)\|_\infty + h \left( L_{FF} \|\Delta y_{F,n+(l+1)/m}\| \right. \\ &\quad \left. + L_{FS} \|\Delta y_{S,n+(l+1)/m}\| + L_{Fz} \|\Delta z_{S,n+(l+1)/m}\| \right). \end{aligned}$$

Summing all micro-steps ( $l = 0, 1, \dots, m-1$ ), using exact IVs at  $t = t_n$  (14) and the bound of the local truncation error  $\delta_{n+l/m}$  (for the implicit Euler)

$$\|\delta_{n+l/m}\| \leq \frac{h^2}{2} \max_{\tau \in [t_{n+l/m}, t_{n+(l+1)/m}]} \|\ddot{y}_F(\tau)\|,$$

we arrive at the statement of the lemma.  $\square$

It remains to estimate  $\Delta \bar{y}_{S,n+l/m}$ ,  $\Delta \bar{z}_{S,n+l/m}$  for all  $l = 0, 1, \dots, m-1$ . The following lemma gives a corresponding bound:

**Lemma 2.** *Under the same settings and assumptions as in Lemma 1, the coupling errors can be bounded by*

- a)  $\|\Delta \bar{y}_{S,n+l/m}\| \leq \frac{1}{2}lh^2(m-l)\|\ddot{y}_S(\tau)\| + \frac{l}{m}\|\Delta y_{S,n+1}\|$  for some  $\tau \in [t_n, t_{n+1}]$ ,
- b)  $\|\Delta \bar{z}_{S,n+l/m}\| \leq \frac{1}{2}lh^2(m-l)\|\ddot{z}_S(\tau)\| + \frac{l}{m}\|\Delta z_{S,n+1}\|$  for some  $\tau \in [t_n, t_{n+1}]$  if  $\bar{z}_{S,n+l/m}$  is achieved by linear interpolation (9),
- c)  $\|\Delta \bar{z}_{S,n+l/m}\| \leq L_{GF}\|\Delta y_{F,n+l/m}\| + L_{GS}\|\Delta \bar{y}_{S,n+l/m}\|$  if the formulation based on the algebraic constraint (10) is used.

*Proof.* a) It holds:

$$\begin{aligned} \Delta \bar{y}_{S,n+l/m} &= y_S(t_{n+l/m}) - \left( \frac{m-l}{m} y_{S,n} + \frac{l}{m} y_{S,n+1} \right) \\ &= y_S(t_{n+l/m}) - \left( \frac{m-l}{m} y_{S,n} + \frac{l}{m} y_S(t_{n+1}) \right) - \frac{l}{m} \Delta y_{S,n+1}. \end{aligned}$$

Then, an error estimation for linear interpolation yields a).

- b) Analogous to a).  
c) We have

$$\|\Delta \bar{z}_{S,n+l/m}\| = \|G(y_{F,n+l/m}, \bar{y}_{S,n+l/m}) - G(y_F(t_{n+l/m}), y_S(t_{n+l/m}))\|.$$

Applying the mean value theorem and using the Lipschitz condition for  $G$  (16), we obtain  $\square$

To estimate  $\Delta y_{F,n+1}$  in terms of  $\Delta y_{S,n+1}$  and  $\Delta z_{S,n+1}$ , we combine the previous lemmas and have as direct consequence:

**Proposition 1.** *Under the same settings and assumptions as in Lemma 1, the error  $\Delta y_{F,n+1}$  can be bounded (using linear interpolation for  $\bar{y}_S$ ):*

- i) for  $\bar{z}_S$  obtained by linear interpolation (9)

$$\begin{aligned} \|\Delta y_{F,n+1}\| \leq C \cdot & \left[ \frac{H^2}{2m} \|\ddot{y}_F(t)\|_\infty + \frac{L_{FS}}{2}(H+h) \left( \frac{H^2}{6} \|\ddot{y}_S(\tau)\|_\infty + \|\Delta y_{S,n+1}\| \right) \right. \\ & \left. + \frac{L_{FZ}}{2}(H+h) \left( \frac{H^2}{6} \|\ddot{z}_S(t)\|_\infty + \|\Delta z_{S,n+1}\| \right) \right]; \end{aligned}$$

- ii) for  $\bar{z}_S$  computed by the non-linear equation (10) and  $h$  restricted to  $0 < 1 - h(L_{FF} - L_{FZ}L_{GF}) < 1$ , then we have the bound

$$\begin{aligned} \|\Delta y_{F,n+1}\| \leq D \cdot & \left[ \frac{H^2}{2m} \|\ddot{y}_F(t)\|_\infty \right. \\ & \left. + \frac{L_{FS} - L_{FZ}L_{GS}}{2}(H+h) \left( \frac{H^2}{6} \|\ddot{y}_S(t)\|_\infty + \|\Delta y_{S,n+1}\| \right) \right] \end{aligned}$$

with constant  $D > \left( \frac{1}{1 - h(L_{FF} - L_{FZ}L_{GF})} \right)^m > 0$ .

Next, we provide estimations for  $\|\Delta y_{S,n+1}\|$  and  $\|\Delta z_{S,n+1}\|$ . We present the result for each coupling approach in a separate subsection.

### 3.3. Accuracy of the Slow Components: Decoupled-Slowest-First

The derivation of an error bound for the slow components is done in two steps: we start with an estimation for the algebraic variables, then the slow differential variables are estimated.

**Lemma 3.** *Let be given an index-1 DAE-IVP (4) fulfilling Ass. 1. Let the approximation  $y_{S,n+1}$ ,  $z_{S,n+1}$  is computed by the *mrIRK1-DAE* scheme (6-8) with macro-step size  $H$  with constant extrapolation for the coupling term  $\bar{y}_{F,n+1} = y_{F,n}$ . Then the error in  $z_S$  can be bounded by*

$$\|\Delta z_{S,n+1}\| \leq H \cdot L_{GF} \|\dot{y}_F(\tau)\| + L_{GS} \|\Delta y_{S,n+1}\| \quad (18)$$

with  $\tau \in [t_n, t_n + H]$  and Lipschitz constants  $L_{GF}, L_{GS}$ .

*Proof.* Solving the algebraic constraint (5), we can write for the local error

$$\Delta z_{S,n+1} = G(y_{F,n}, y_{S,n+1}) - G(y_F(t_{n+1}), y_S(t_{n+1})).$$

Applying the mean value theorem, Lipschitz continuity of  $G$  and norms (similar to Lemma 1), we obtain

$$\|\Delta z_{S,n+1}\| \leq L_{GF} \|y_{F,n} - y_F(t_{n+1})\| + L_{GS} \|\Delta y_{S,n+1}\|.$$

Then using  $y_{F,n} = y_F(t_n)$  and the mean value theorem, the proof is completed.  $\square$

Next, we estimate the error in the  $y_S$ .

**Proposition 2.** *Under the same settings and assumptions as in Lemma 3 and a restricted macro-step size  $H$ , such that  $0 < 1 - H(L_{SS} + L_{SZ}L_{GS}) < 1$  holds, the error in  $y_S$  is bounded by*

$$\|\Delta y_{S,n+1}\| \leq \frac{H^2}{1 - H(L_{SS} + L_{SZ}L_{GS})} \left[ (L_{SS} + L_{SZ}L_{GS}) \|\dot{y}_S(t)\|_\infty + 2L_{SZ}L_{GF} \|\dot{y}_F(t)\|_\infty + \frac{1}{2} \|\ddot{y}_S(t)\|_\infty \right]. \quad (19)$$

*Proof.* By Taylor expansion of  $y_S(t_{n+1})$  with expansion point  $t_n$ , we obtain

$$\Delta y_{S,n+1} = H[f_S(y_{F,n}, y_{S,n+1}, z_{S,n+1}) - f_S(y_{F,n}, y_{S,n}, z_{S,n})] - \frac{H^2}{2} \ddot{y}_S(\tau)$$

for some  $\tau \in [t_n, t_n + H]$ . Applying norms and Lipschitz continuity, we get

$$\|\Delta y_{S,n+1}\| \leq H \left[ L_{SS} \|\Delta y_{S,n+1}\| + L_{SS} \|y_S(t_{n+1}) - y_S(t_n)\| + L_{SZ} \|\Delta z_{S,n+1}\| + L_{SZ} \|z_S(t_{n+1}) - z_S(t_n)\| \right] + \frac{H^2}{2} \|\ddot{y}_S(t)\|_\infty.$$

Again, mean value theorem and Lemma 3 lead to

$$\|\Delta y_{S,n+1}\| \leq H \left[ L_{SS} \|\Delta y_{S,n+1}\| + HL_{SS} \|\dot{y}_S(\theta)\| + HL_{SZ}L_{GF} \|\dot{y}_F(\xi)\| + L_{SZ}L_{GS} \|\Delta y_{S,n+1}\| + HL_{SZ} \|\dot{z}(\rho)\| \right] + \frac{H^2}{2} \|\ddot{y}_S(t)\|_\infty$$

for  $\theta, \xi, \rho \in [t_n, t_n + H]$ . By using Ass. 1(ii), we have  $\|\dot{z}_S(t)\| \leq L_{GF} \|\dot{y}_F(t)\| + L_{GS} \|\dot{y}_S(t)\|$ . Inserting this, we can finally solve for  $\|\Delta y_{S,n+1}\|$ .  $\square$

Summing up, we have:

**Corollary 1.** *Under the same settings and assumptions as in Prop. 2, the decoupled-slowest-first, mrIRK1-DAE method (6-8) has consistency order 1 in the differential variables and the error in the algebraic variables is  $\mathcal{O}(H)$  (under the above step size restrictions).*

### 3.4. Accuracy of the Slow Components: Coupled-Slowest-First

Here,  $y_{S,n+1}$  and  $z_{S,n+1}$  depend on the auxiliary variable  $y_{F,n+1}^*$ . The next two lemmas give estimates for the algebraic and differential variables:

**Lemma 4.** *We consider an index-1 DAE-IVP (4) fulfilling the Ass. 1. We apply the coupled-slowest-first (11), mrIRK1-DAE method. Then, the error in the slow changing, algebraic variable can be estimated by*

$$\|\Delta z_{S,n+1}\| \leq L_{GF} \|\Delta y_{F,n+1}^*\| + L_{GS} \|\Delta y_{S,n+1}\|.$$

The proof is similar to the deduction of Lemma 3.

**Lemma 5.** *Under the same settings and assumptions as in Lemma 4, the error in the differential variables in the coupled-slowest-first approach can be bounded as follows:*

$$M(H, H) \left( \frac{\|\Delta y_{F,n+1}^*\|}{\|\Delta y_{S,n+1}\|} \right) \leq \left( \frac{\frac{H^2}{2} \|\ddot{y}_F(t)\|_\infty}{\frac{H^2}{2} \|\ddot{y}_S(t)\|_\infty} \right), \quad (20)$$

$$\text{with } M(H_1, H_2) := \begin{pmatrix} 1 - H_1(L_{FF} + L_{FZ}L_{GF}) & -H_1(L_{FS} + L_{FZ}L_{GS}) \\ -H_2(L_{SF} + L_{SZ}L_{GF}) & 1 - H_2(L_{SS} + L_{SZ}L_{GS}) \end{pmatrix}$$

The inequality in (20) has to be understood componentwise.

*Proof.* For  $\Delta y_{S,n+1}$ , we add  $\pm [y_S(t_n) - H f_S(y_F(t_{n+1}), y_S(t_{n+1}), z_S(t_{n+1}))]$ . By the mean value theorem, we deduce

$$\begin{aligned} \|\Delta y_{S,n+1}\| &\leq H \left[ L_{SF} \|\Delta y_{F,n+1}^*\| + L_{SS} \|\Delta y_{S,n+1}\| + L_{SZ} \|\Delta z_{S,n+1}\| \right] \\ &\quad + \left\| \int_0^H \tau \dot{y}(t_n + \tau) d\tau \right\|. \end{aligned}$$

Employing Lemma 4 for  $\|\Delta z_{S,n+1}\|$  and again the mean value theorem, we find

$$[1 - H(L_{SS} + L_{SZ}L_{GS})] \|\Delta y_{S,n+1}\| - H(L_{SF} + L_{SZ}L_{GF}) \|\Delta y_{F,n+1}^*\| \leq \frac{H^2}{2} \|\ddot{y}_S(t)\|_\infty.$$

Analogously, one can deduce the estimate for  $\Delta y_{F,n+1}^*$ .  $\square$

To solve the estimate (20) for the error in the differential variables, we need that  $M(H_1, H_2)$  is an M-matrix in  $\mathbb{R}^{2 \times 2}$  (later we will need this more general version). In fact, for  $H_1, H_2 > 0$  small enough, the diagonal entries are positive (off-diagonals are always negative). Thus, we have

**Proposition 3.** *Let the same settings and assumptions apply as in Lemma 4 (coupled-slowest-first). And the step-size  $H$  be restricted such that holds:*

$$H(L_{FF} + L_{FZ}L_{GF}) < 1 \quad \text{and} \quad H(L_{SS} + L_{SZ}L_{GS}) < 1. \quad (21)$$

Then we have

$$\begin{aligned} \|\Delta y_{S,n+1}\| &\leq \frac{1}{\det(M(H,H))} \left[ \frac{H^3}{2} (L_{SF} + L_{SZ}L_{GF}) \|\ddot{y}_F(t)\|_\infty \right. \\ &\quad \left. + \frac{H^2}{2} (1 - H(L_{FF} + L_{FZ}L_{GF})) \|\ddot{y}_S(t)\|_\infty \right] \\ \|\Delta y_{F,n+1}^*\| &\leq \frac{1}{\det(M(H,H))} \left[ \frac{H^3}{2} (L_{FS} + L_{FZ}L_{GS}) \|\ddot{y}_S(t)\|_\infty \right. \\ &\quad \left. + \frac{H^2}{2} (1 - H(L_{SS} + L_{SZ}L_{GS})) \|\ddot{y}_F(\tau)\|_\infty \right]. \end{aligned}$$

The last results give the consistency:

**Corollary 2.** *Under the same settings and assumptions as in Prop. 3, the coupled-slowest-first, mrIRK1-DAE method (6-8) has consistency order 1 when applied to semi-explicit DAEs of index-1.*

### 3.5. Accuracy in Compound Step (Coupled-First-Step)

We consider (12) and address first the algebraic variable and then the dynamic variables:

**Lemma 6.** *We consider the DAE-IVP (4) fulfilling Ass. 1 and apply the coupled-first-step (12), mrIRK1-DAE method. We find for  $\|\Delta z_{S,n+1}\|$ :*

$$\|\Delta z_{S,n+1}\| \leq L_{GF} \|\Delta y_{F,n+1}\| + L_{GS} \|\Delta y_{S,n+1}\| + HL_{GF} \|\dot{y}_F(t)\|_\infty. \quad (22)$$

*Proof.* Using the implicit function for  $z_S$  (5), we can estimate (cf. Lemma 3)

$$\|\Delta z_{S,n+1}\| \leq L_{GS} \|\Delta y_{S,n+1}\| + L_{GF} \|y_{F,n+1/m} - y_F(t_{n+1})\|.$$

In the second summand we add and subtract  $y_F(t_{n+1/m})$ . Applying the mean value theorem leads to the statement of the lemma.  $\square$

Similar as the deduction of Lemma 5, we can obtain:

**Lemma 7.** *Under the same settings and assumptions as in Lemma 6, the error in the differential variables (coupled-first-step) can be estimated as*

$$M(h, H) \begin{pmatrix} \|\Delta y_{F, n+1/m}\| \\ \|\Delta y_{S, n+1}\| \end{pmatrix} \leq \begin{pmatrix} R_F \\ R_S \end{pmatrix}$$

with  $M(h, H)$  given in (20) and

$$\begin{aligned} R_F &= \frac{H^2}{m} (L_{FS} + L_{FZ}L_{GS}) \|\dot{y}_S(t)\|_\infty + \frac{2H^2}{m} (L_{FZ}L_{GF}) \|\dot{y}_F(t)\|_\infty + \frac{h^2}{2} \|\ddot{y}_F(t)\|_\infty, \\ R_S &= H^2 (L_{SF} + L_{SZ}L_{GF}) \|\dot{y}_F(t)\|_\infty + \frac{H^2}{2} \|\ddot{y}_S(t)\|_\infty. \end{aligned}$$

Again, the M-matrix property of  $M(h, H)$  ( $h, H$  small enough) leads to: (cf. Prop. 3)

**Proposition 4.** *Under the same settings and assumptions as in Lemma 6 and step-size restrictions*

$$H(L_{SS} + L_{SZ}L_{GS}) < 1 \quad \text{and} \quad h(L_{FF} + L_{FZ}L_{GF}) < 1, \quad (23)$$

the coupled-first-step, *mrIRK1-DAE* scheme applied DAE-IVP (4) is of consistency order 1 in the differential variables  $y_F$  and  $y_S$ . The error in the slow changing, algebraic variable is in  $\mathcal{O}(H)$ .

### 3.6. Summary

We conclude for Section 3:

**Theorem 1.** *For all versions of the mrIRK1-DAE method applied to the DAE-IVP (4) the differential variables ( $y_F, y_S$ ) have consistency order 1. The algebraic variable ( $z_S$ ) reach order 1 only in the coupled-slowest-first approach. For the other coupling approaches, the error  $\|\Delta z_S\|$  is always in  $\mathcal{O}(H)$ . Under the additional assumption*

$$\frac{\partial}{\partial y_F} G(y_F, y_S) = 0 \quad (24)$$

for  $t \in [t_0, t_{end}]$  we have order 1 also in the algebraic variable ( $z_S$ ) in all coupling approaches.

*Proof.* It only remains to show order 1 in  $z_S$  for *Decoupled-Slowest-First* and *Coupled-First-Step*: Since (24), we have  $L_{GF} = 0$  in (18) and (22) and we end up with

$$\|\Delta z_S\| = \mathcal{O}(\|\Delta y_S\|).$$

□

*Remark:* The slow changing variables ( $y_S, z_S$ ) of a multirate DAE-IVP depends only weakly on  $y_F$ , therefore  $\left\| \frac{\partial}{\partial y_F} G(y_F, y_S) \right\|$  is small and can be neglected in most cases.

Next, it is shown that the reduced consistency order in the algebraic variable does not influence the convergence of the scheme.

#### 4. Convergence

Now, we investigate the error propagation over several macro-steps. For the index-1 DAE-IVP (4),  $(y_{F,n}, y_{S,n}, z_{S,n})$  denotes the mrIRK1-DAE approximation at  $t_n$  after  $n$  macro-steps. For any components  $x = x(t)$  of the unknowns, the global error reads

$$e(x, t_n) := x_n - x(t_n).$$

We show that  $e(y_F, t_n)$ ,  $e(y_S, t_n)$ ,  $e(z_S, t_n)$  are in  $\mathcal{O}(H)$ . To this end, we recall the following theorem from [11]: given a semi-explicit DAE-IVP of index-1 (4), we apply a general one-step method

$$\begin{aligned} y_{k+1} &= y_k + \hat{h} \cdot \Phi(y_k, z_k, \hat{h}), \\ z_{k+1} &= \Psi(y_k, z_k, \hat{h}) \end{aligned}$$

with  $y^\top = (y_F^\top, y_S^\top)$ , a constant step size  $\hat{h}$ , a differential update function  $\Phi$  and an algebraic update function  $\Psi$ . We remark that  $\Phi$  and  $\Psi$  are only formally explicit. If the method has consistency order  $p$  for the differential variables  $y$ , as well as  $p - 1$  for algebraic variables  $z$  and if the algebraic update function satisfies the following perturbation condition

$$\left\| \frac{\partial \Psi(y, z, 0)}{\partial z} \right\| \leq \alpha < 1 \quad (25)$$

in a neighbourhood of the solution, then the one-step method has convergence order  $p$ .

For  $p = 1$  this statement holds for the mrIRK1-DAE method:

**Theorem 2.** *We apply the mrIRK1-DAE method to the index-1 DAE-IVP (4) fulfilling Ass. 1. We may choose any coupling variant: coupled-slowest-first, decoupled-slowest-first, coupled-first-step.  $H$  and  $m$  are chosen such that (21) and (23) are fulfilled. Then we get for the global error*

$$e(y_F, t_n) = \mathcal{O}(H), \quad e(y_S, t_n) = \mathcal{O}(H), \quad e(z_S, t_n) = \mathcal{O}(H).$$

*Proof.* We check the assumptions of the theorem from [11] (mentioned above):

*One-Step Method.* All discussed formulations of the mrIRK1-DAE scheme define the approximations  $y_{F,n+1}, y_{S,n+1}$  and  $z_{S,n+1}$  at  $t_{n+1}$  after one macro step as functions of the approximations  $y_{F,n}, y_{S,n}$  and  $z_{S,n}$  at  $t_n$ .

*Consistency.* Theorem 1 showed that we have consistency order 1 for the differential variables and at least order  $\mathcal{O}(H)$  for the algebraic variables (for any variant).

*Perturbation Condition (25).* We discuss the *coupled-first-step* approach. Using (5), we have  $z_{n+1} = G(y_{F,n+1}^*, y_{S,n+1})$ . Inserting  $y_{F,n+1}^*$  and  $y_{S,n+1}$ , we get

$$z_{n+1} = G(y_{F,n} + H f_F(y_{F,n+1}^*, y_{S,n+1}, z_{S,n+1}), y_{S,n} + H f_S(y_{F,n+1}^*, y_{S,n+1}, z_{S,n+1})).$$

Hence  $z_{n+1}$  does not depend on  $z_n$  and the estimate on  $\Psi$  is fulfilled. In a similar way, we can deduce this result for the other coupling approaches.  $\square$

The following numerical simulations confirm this analytical result.

## 5. Numerical Results

For the numerical verification, we consider two DAE-systems.

### 5.1. Extended Prothero-Robinson Equation

An extended Prothero-Robinson test equation for semi-explicit DAEs [12] reads in our settings as follows

$$\begin{pmatrix} \dot{y} \\ 0 \end{pmatrix} = \begin{pmatrix} A - BF & B \\ C - DF & D \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} + \begin{pmatrix} -A\eta(t) - B\zeta(t) + \dot{\eta}(t) \\ -C\eta(t) - D\zeta(t) \end{pmatrix} \quad (26)$$

with  $y(t) = (y_S(t), y_F(t))^\top \in \mathbb{R}^2$  and  $z(t) = z_S(t) = (z_{S1}(t), z_{S2}(t))^\top \in \mathbb{R}^2$  and given functions  $\eta$  and  $\zeta$ . For the simulation we choose the following data:

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\eta(t) = (\sin(2\pi 10^6 t), 2 \cos(2\pi 10^7 t))^\top, \quad \zeta(t) = (2 \cos(t), 7t)^\top.$$

Since  $D$  is regular (26) is of index-1 and consistent initial values are given by  $(y_S(0), y_F(0), z_{S1}(0), z_{S2}(0))^\top = (0, 2, 2, 0)^\top$ . Notice that the solution of (26) is

$$y(t) = \eta(t), \quad z(t) = F\eta(t) + \zeta(t).$$

We apply the mrIRK1-DAE method to the DAE (26) on  $[t_0, t_{\text{end}}] = [0, 10^{-6}s]$  using all three coupling approaches. We use different macro-step sizes  $H = 2^{2-i} \cdot 10^{-8}$  for  $i = 0, \dots, 7$ , multirate factor  $m = 10$  and  $m = 20$ . We investigate the absolute value of  $e(x, t_{\text{end}})$  for all four components, that is the algebraic components are treated separately.

Fig. 1 shows the convergence order for the *decoupled-slowest-first* strategy. We observe order 1 apart from  $z_{S2}$ , where the simulation indicates order 2, see Fig. 1d). This phenomenon is caused by the coupling structure and data in the DAE (26).

Fig. 2 gives the simulation results of the *coupled-slowest-first* strategy for the differential variables, which are quite similar to the *decoupled-slowest-first* case (Fig. 1). The convergence of the algebraic variable are the same in both coupling approaches so no figures are given.

Simulation results for the *coupled-first-step* strategy are given in Fig. 3. The behaviour of the slow variables is the same as for the other coupling approaches, again we skip figures for the algebraic variable.

The fast variable  $y_F$  shows overall convergence, but its behaviour is slightly more irregular than the others (Fig. 3a). Moreover, we compare with a higher multirate factor of  $m = 20$ . The *decoupled-slowest-first* and *coupled-slowest-first* show the same convergence properties as for  $m = 10$  (no figures given).

For the *coupled-first-step* case with  $m = 20$ , we observe order 1, see Fig. 4. We remark that only in this coupling approach the consistency on the macro-step depends on both  $H$ ,  $h$  and thus on  $m$  (23).

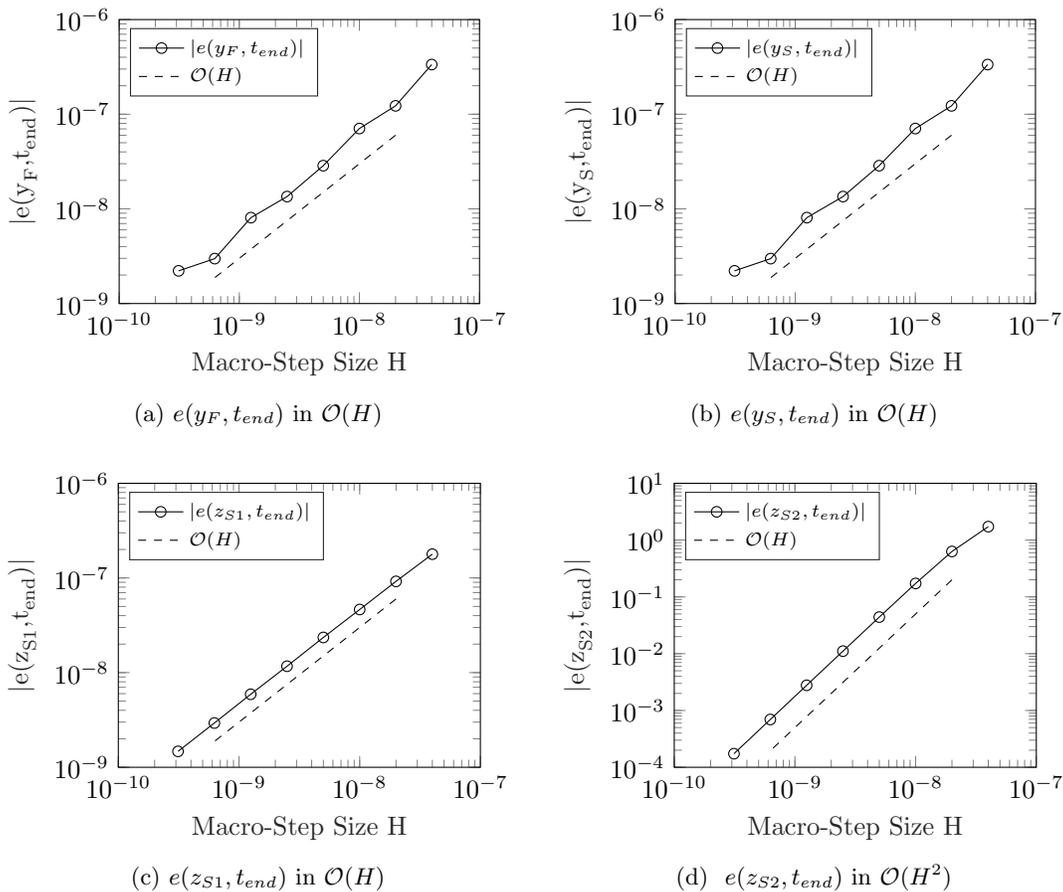


Figure 1: Order of convergence for the *decoupled-slowest-first* approach (m=10): a)-c) order 1, d) order 2.

### 5.2. Field-Circuit Coupled System

We consider a field-circuit coupled system, a circuit diagram is given in Figure 5, for details see [13]. The system equations are given by

$$C\dot{e}_1(t) = G(e_1(t) - U_{in}(t)) - I_{CO}(t) \tag{27}$$

$$E\dot{x}_S(t) = Ax_S(t) + Be_1(t). \tag{28}$$

The fast-changing subsystem (27) describes a node potential  $e_1$  in an electrical circuit with capacitance  $C = 1\text{nF}$ , conductance  $G = 0.01\text{S}$ , input voltage  $U_{in}(t) = 45.5 \cdot 10^3 \sin(900\pi t) + 10^3 \sin(45000\pi t)$  and coupling current  $I_{CO}$ . The slow-changing subsystem (28) results from a finite-element discretisation of a magneto-quasistatic equation, which describes the electric field of a 2D-transformer with state space vector  $x_S$ , system matrix  $A$ , input matrix  $B$  and mass matrix





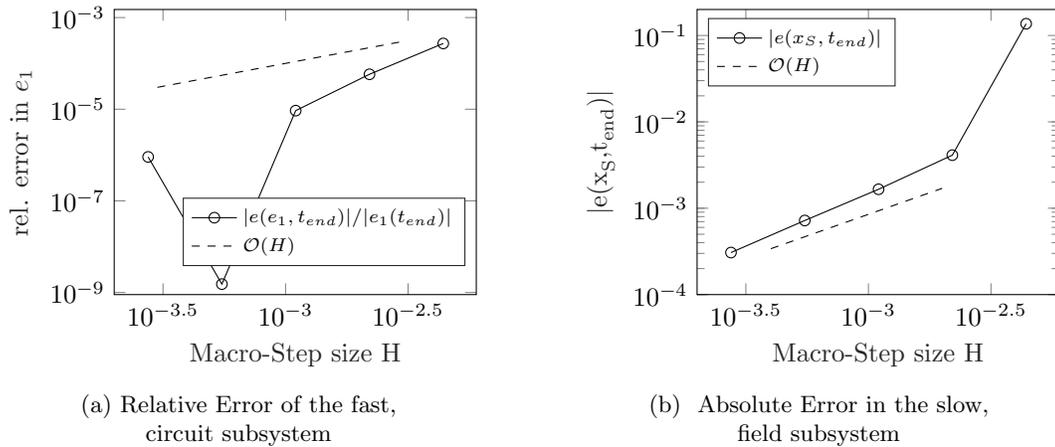


Figure 6: Convergence order for the *coupled-fastest-first* approach for the field-circuit coupled system

- [3] A. Kværnø, P. Rentrop, Low order multirate Runge-Kutta methods in electric circuit simulation, preprint No. 2/99, NTNU Trondheim (1999).
- [4] W. Hundsdorfer, V. Savcenco, Analysis of a multirate theta-method for stiff ODEs, *Applied Numerical Mathematics* 59 (2009) 693 – 706.
- [5] M. Günther, A. Sandu, Multirate generalized additive Runge Kutta methods, *Numerische Mathematik* 133 (3) (2016) 497–524.
- [6] K. E. Brenan, S. L. Campbell, L. R. Petzold, *Numerical solution of initial-value problems in differential-algebraic equations*, Society for Industrial and Applied Mathematics Philadelphia, 1995.
- [7] E. Hairer, G. Wanner, *Solving ordinary differential equations II: stiff and differential-algebraic problems*, Springer Berlin Heidelberg, 2002.
- [8] P. Kunkel, V. Mehrmann, *Differential-algebraic equations: analysis and numerical solution*, EMS textbooks in mathematics, European Mathematical Society Zürich, 2006.
- [9] M. Striebel, Hierarchical mixed multirating for distributed integration of DAE network equations in chip design, Ph.D. thesis, University of Wuppertal. *Fortschritt-Berichte VDI, Reihe 20, Nr. 404*, VDI-Verlag Düsseldorf, 2006.
- [10] A. Verhoeven, Redundancy reduction of IC models by multirate time-integration and model order reduction, Ph.D. thesis, Technische Universiteit Eindhoven (2008).
- [11] P. Deuffhard, E. Hairer, J. Zugck, One-step and extrapolation methods for differential-algebraic systems, *Numerische Mathematik* 51 (5) (1987) 501–516.
- [12] A. Bartel, M. Brunk, S. Schöps, On the convergence rate of dynamic iteration for coupled problems with multiple subsystems, *Journal of Computational and Applied Mathematics* 262 (2014) 14–24.

- [13] C. Hachtel, A. Bartel, M. Günther, J. Kerler-Back, T. Stykel, Multirate dae/ode-simulation and model order reduction for coupled field-circuit systems, in: W. Langer, Ulrich and Amrhein, W. Zulehner (Eds.), *Scientific Computing in Electrical Engineering at SCEE 2016*, St. Wolfgang, Austria, October 2016, Springer, Berlin, 2018.
- [14] J. Kerler-Back, T. Stykel, Model reduction for linear and nonlinear magnetoquasistatic equations, *International Journal for Numerical Methods in Engineering* 111 (13) (2017) 1274–1299.